Python Assignment

Bhupender Kumar
24.01.2022

Problem Statement

· Imagine you are working as a Data Scientist for an Online Wine Shop named “The Wine Land”

· As the name suggests, the online store specializes in selling different varieties of wines.

· The online store receives a decent amount of traffic and reviews from its users.

Leverage the “reviews” data and draw actionable insights from it.

Objective

- Derive the top 5 actionable Insights from the Data.
- Build a predictive model for predicting the wine “variety”. Provide the output along with all features to a CSV file. Both Training & test data is provided here
- Submit the source code used for building models in a zip or share the link to the GitHub repository.
- Also submit a short summary: Model used, features extracted, Model accuracy in train data set, along with some visualization of data and findings.

Insights

1. Price range of the wine is between 4 - 3300 currency. 50% of the wine price is below 27 currency.
2. Ratings obtained maximum time are 87(10306 times) and 88(10504 times).
3. Correlation between Points and price is 40. This states there is positive but not very good relationship between review points and price of wine.
4. Most reviewed wine is “Pinot Noir” 10587 times.
5. Top producer of wine is Testarossa winery.
6. Top wine producer country is USA.

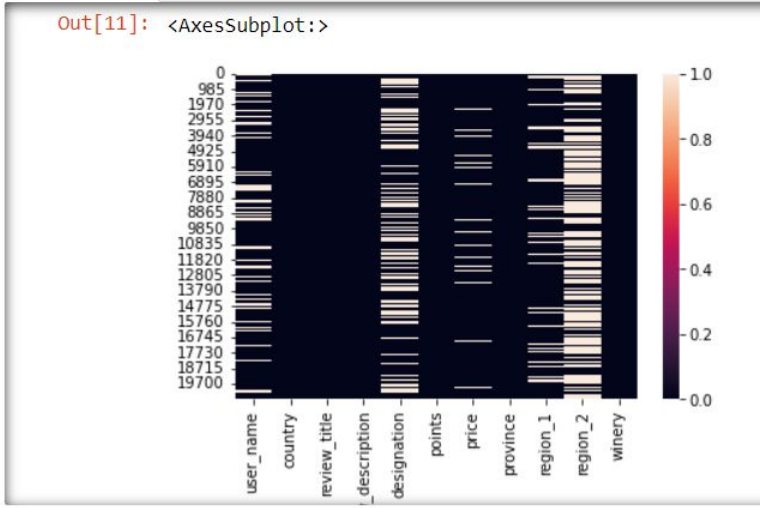
Summary

Model used	Random Forest Logistics Regression
Features extracted	TF-IDF → It is used to extract the Numpy matrix from the text and used maximum features 5000 for model building.
Model accuracy in train data set	Random Forest accuracy → 93% Logistics Regression accuracy→ 94%
Visualization of data and findings	Seaborn → Heatmap for missing value in data set and correlation between points and price of wine.

Visualization

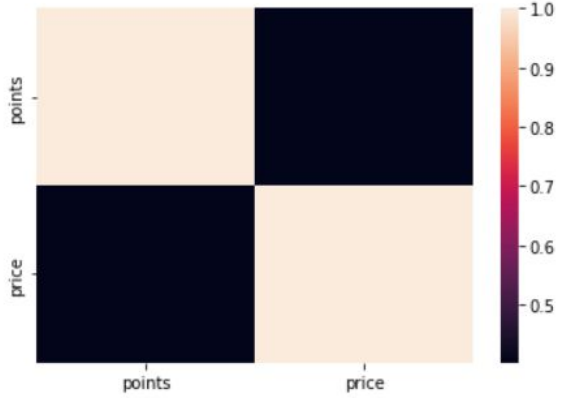
```
In [11]: sns.heatmap(test.isnull())
```

```
Out[11]: <AxesSubplot:>
```

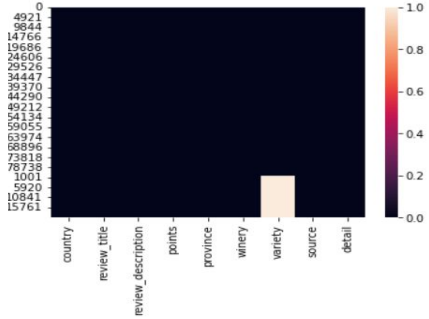



```
In [25]: sns.heatmap(corr)
```

```
Out[25]: <AxesSubplot:>
```




```
ns.heatmap(winedata_new.isnull())
AxesSubplot:>
```



Thank You